

Fluctuations of fitness distributions and the rate of Muller's ratchet

Richard A. Neher

Max Planck Institute for Developmental Biology, Tübingen, 72076, Germany

Boris I. Shraiman

Kavli Institute for Theoretical Physics and Department of Physics, University of California, Santa Barbara, CA 93106

(Dated: June 29, 2012)

The accumulation of deleterious mutations is driven by rare fluctuations which lead to the loss of all mutation free individuals, a process known as Muller's ratchet. Even though Muller's ratchet is a paradigmatic process in population genetics, a quantitative understanding of its rate is still lacking. The difficulty lies in the nontrivial nature of fluctuations in the fitness distribution which control the rate of extinction of the fittest genotype. We address this problem using the simple but classic model of mutation selection balance with deleterious mutations all having the same effect on fitness. We show analytically how fluctuations among the fittest individuals propagate to individuals of lower fitness and have a dramatically amplified effects on the bulk of the population at a later time. If a reduction in the size of the fittest class reduces the mean fitness only after a delay, selection opposing this reduction is also delayed. This delayed restoring force speeds up Muller's ratchet. We show how the delayed response can be accounted for using a path integral formulation of the stochastic dynamics and provide an expression for the rate of the ratchet that is accurate across a broad range of parameters.

By weeding out deleterious mutations, purifying selection acts to preserve a functional genome. In sufficiently small populations, however, weakly deleterious mutations can by chance fix. This phenomenon, termed Muller's ratchet (FELSENSTEIN, 1974; MULLER, 1964), is especially important in the absence of recombination and is thought to account for the degeneration of Y-chromosomes (RICE, 1987) and for the absence of long lived asexual lineages (LYNCH *et al.*, 1993).

A click of Muller's ratchet refers to the loss of the class of individuals with the smallest number of deleterious mutations. To understand the processes responsible for such a click, it is useful to consider a simple model of accumulation of deleterious mutations with identical effect sizes illustrated in Fig. 1. Because of mutations, the population spreads out along the fitness axis, which in this model is equivalent to the number of deleterious mutations in a genome. The population can hence be grouped into discrete classes each characterized by the number of deleterious mutations. Mutation carries individuals from classes with fewer to classes with more mutations, hence shifting the population to the left. This tendency is opposed by selection, which amplifies fit individuals on the right, while decreasing the number of unfit individuals on the left. These opposing trends lead to a steady balance, at least in sufficiently large populations. However, in addition to selection and mutation, the distribution of individuals among fitness classes is affected by fluctuations in the number of offspring produced by individuals of different classes, i.e., by genetic drift. Such fluctuations are stronger (in relative terms) in smaller populations and in particular in classes which carry only a small number of individuals. When mutation rate is high and selection is weak, the class of individuals with the smallest number of mutations ($k = 0$ in Fig. 1) contains only few individuals and is therefore susceptible to accidental extinction - an event that corresponds to the "click" of Muller's ratchet.

Despite the simplicity of the classic model described above, understanding the rate of the ratchet has been a challenge and remains incomplete (ETHERIDGE *et al.*, 2007; GESSLER, 1995; GORDO and CHARLESWORTH, 2000; HIGGS and WOODCOCK, 1995; JAIN, 2008; STEPHAN *et al.*, 1993; STEPHAN and KIM, 2002; WAXMAN and LOEWE, 2010). Here, we revisit this problem starting with the systematic analysis of fluctuations in the distribution of the population among different fitness classes. We show that fitness classes do not fluctuate independently. Instead, there are collective modes affecting the entire distribution which relax on different time scales. Having identified these modes, we calculate the fluctuations of the number of individuals in the fittest class and show how these fluctuations affect the mean fitness. Fluctuations in mean fitness feed back on the fittest class with a delay and thereby control the probability of extinction. These insights allow us to arrive at a better approximation to the rate of the ratchet. In particular, we show that the parameter introduced in the earlier work (GORDO and CHARLESWORTH, 2000; HAIGH, 1978; STEPHAN *et al.*, 1993) to parameterize the effective strength of selection in the least loaded class is not a constant but depends on the ratio of the mutation rate and the effect size of mutations. We use the path integral representation of stochastic processes, borrowed from physics (FEYNMAN and HIBBS, 1965) to describe the dynamics of the fittest class and arrive at an approximation of the rate of Muller's ratchet that is accurate across a large parameter range.

Understanding the rate of the ratchet is important, for example to estimate the number of beneficial mutations

required to halt the ratchet and prevent the mutational melt-down of a population (GOYAL *et al.*, 2012; LYNCH *et al.*, 1993; PFAFFELHUBER *et al.*, 2012) (for an in-depth and up-to-date discussion of the importance of deleterious mutations we refer the reader to CHARLESWORTH (2012)). Furthermore, fluctuations of fitness distributions are a general phenomenon with profound implications for the dynamics of adaptation and genetic diversity of populations. Below we shall place our approach into the context of the recent studies of the dynamics of adaptation in populations with extensive non-neutral genetic diversity (DESAI and FISHER, 2007; NEHER *et al.*, 2010; ROUZINE *et al.*, 2003; TSIMRING *et al.*, 1996). The study of fluctuations in the approximately stationary state of mutation selection balance which we present here is a step towards more general quantitative theory of fitness fluctuations in adapting populations.

I. MODEL AND METHODS

We assume that mutations happen at rate u and that each mutation reduces the growth rate of the genotype by $s \ll 1$. Within this model, proposed and formalized by HAIGH, individuals can be categorized by the number of deleterious mutations they carry. The equation describing the fitness distribution in the population, i.e., what part n_k of the population carries k deleterious mutations, is given by

$$\frac{d}{dt}n_k = s(\bar{k} - k)n_k - un_k + un_{k-1} + \sqrt{n_k}\eta_k \quad (1)$$

where $\bar{k} = N^{-1} \sum_k kn_k$ ($\sum_k n_k = N$) and the last term accounts for fluctuations due to finite populations, i.e., genetic drift, and has the properties of uncorrelated Gaussian white noise with $\langle \eta_k(t)\eta_l(t') \rangle = \delta_{kl}\delta(t-t')$. In the infinite population limit, this equation has the well known steady state solution $\bar{n}_k = Ne^{-\lambda}\lambda^k/k!$, where $\lambda = u/s$. A time dependent analytic solution of the deterministic model has been described in (ETHERIDGE *et al.*, 2007).

Note that we have deviated slightly from the standard model, which assumes that genetic drift amounts to a binomial resampling of the distribution with the current frequencies $N^{-1}n_k$. This choice would result in off-diagonal correlations between noise terms that stem from the constraint that the total population size is strictly constant. This exact population size constraint is an arbitrary model choice which we have relaxed to simplify the algebra. Instead, we control the population size by a soft constraint which keeps the population constant on average but allows small fluctuations of N . The implementation of this constraint is described explicitly below. We confirmed the equivalence of the two models by simulation.

Computer simulations

We implemented the model as a computer simulation with discrete generations, where each generation is produced by a weighted resampling of the previous generation. Specifically,

$$n_k(t+1) \leftarrow \text{Poisson} \left(\frac{1}{\bar{W}} \sum_{i=0}^k \frac{e^{-u}u^i}{i!} (1-s)^{k-i} n_{k-i}(t) \right) \quad (2)$$

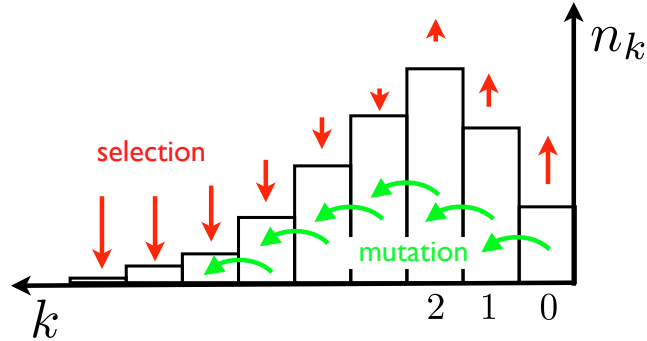


FIG. 1 Deleterious mutation selection balance. The population is distributed among classes of individuals carrying k deleterious mutations. Classes with few mutations grow due to selection (red arrows), but loose individuals through mutations (green arrows), while classes with many mutations are selected against but replenished by mutations.

Symbol	Description
N, u, s	population size, mutation rate, and mutation effect
$n_k(t), \bar{n}_k$	number of individuals in class k at time t , steady state value
$\lambda = u/s, \tau = ts$	rescaled mutation rate and time
$x_k(\tau), \bar{x}_k$	population frequency in class k at time τ , steady state value
z_τ, \bar{z}	abbreviation for $x_0(\tau)$ and \bar{x}_0
\bar{k}	mean fitness: $\bar{k} = \sum_k k x_k$
$\delta x_k, \delta \bar{k}$	deviations from steady state: $\delta x_k = x_k - \bar{x}_k, \delta \bar{k} = \lambda - \bar{k}$
$S_\lambda(\{z_\rho\})$	path integral action depending on the path $\{z_\rho\}$ with $0 \leq \rho \leq \tau$.
$S_\lambda^*(z_\tau, z_0), z_\rho^*$	extremal action and the associated path depending on the endpoints z_τ, z_0
$S_\lambda^*(z)$	long time limit of $S_\lambda^*(z_\tau, z_0)$ with $z = z_\tau$
$P_\tau(z_\tau z_0)$	propagator from z_0 to z_τ in time τ
$p(z)$	steady state distribution of z
γ	rate of the ratchet in units of s
σ^2	variance of x_0 depending on Ns and λ
ζ^2	rescaled variance of x_0 depending on λ only: $\zeta^2 = Ns e^\lambda \sigma^2$
$\psi_k^{(i)}, \phi_k^{(i)}$	k -th component of right and left eigenvectors with eigenvalue $-i$
$a_i(\tau)$	projection of δx_k on $\phi_k^{(i)}$
α	parameter of the effective potential confining x_0 (traditionally $\alpha = 0.5-0.6$)

TABLE I List of symbols

where \bar{W} is the mean fitness $\bar{W} = CN^{-1} \sum_k (1-s)^k n_k$ and $C = \exp(NN_0^{-1} - 1)$ is an adjustment made to the overall growth rate to keep the population size approximately at N_0 . This specific discretization is chosen because it has exactly the same stationary solution as the continuous time version above (HAIGH, 1978). The simulation was implemented in Python using the scientific computing environment SciPy (OLIPHANT, 2007). If the parameter of the Poisson distribution was larger than 10^4 , a Gaussian approximation to the Poisson distribution was used to avoid integer overflows.

To determine the ratchet rate, the population was initialized with its steady state expectation \bar{n}_k , allowed to equilibrate for 10^4 generations, and then run for further $T = 10^8$ generations. Over these 10^8 generations, the number of clicks of the ratchet were recorded and the rate estimated as clicks per generation.

The source code of the programs, along with a short documentation, is available as supplementary material. In addition, we also provide some of the raw data and the analysis scripts producing the figures as they appear in the manuscript.

Numerical determination of the most likely path.

The central quantities in our path integral formulation of the rate of Muller's ratchet are i) the most likely path to extinction and ii) the associated minimal action S_λ^* . To determine the most likely path to extinction of the fittest class, we discretize the trajectory into m equidistant time points ρ_i between 0 and τ , where $x_0(0) = \bar{x}_0$ and $x_0(\rho_m) = 0$. For a given set of $x_0(\rho_i)$, a continuous path $x_0(\rho)$ is generated by linear interpolation. For a given trajectory $x_0(\rho)$, we determine the mean fitness by solving the deterministic equations for $x_k(\rho)$, $k \geq 1$. Note that in this scheme, the only independent variable is the path $x_0(\rho)$, all other degrees of freedom are slaved to $x_0(\rho)$. From $x_0(\rho)$ and the resulting $\bar{k}(\rho)$, we calculate the action $S_\lambda(\{x_0(\rho)\})$ as defined in Eq. (28). $S_\lambda(\{x_0(\rho)\})$ is then minimized by changing the values of $x_0(\rho_i)$, $0 < i < m$, using the simplex minimization algorithm implemented in SciPy (OLIPHANT, 2007). To speed up convergence, the minimization is first done with a small number of pivot points ($m = 4$), which is increased in steps of 2 to $m = 24$. The total time $\tau = 20$ (in units of s^{-1}) was used which is sufficiently large to make the result independent of τ . The code used for the minimization is provided as supplementary material.

II. RESULTS AND DISCUSSION

Fluctuations of the size n_0 of the least loaded class can lead to its extinction. In the absence of beneficial mutations this class is lost forever (MULLER, 1964), and the resulting accumulation of deleterious mutations could have dramatic

evolutionary consequences. Considerable effort has been devoted in understanding this process, and it has been noted that the rate at which the fittest class is lost depends strongly on the average number of individuals in the top class \bar{n}_0 (HAIGH, 1978). Later studies have shown that the rate is exponentially small in $\bar{n}_0 s$ if $\bar{n}_0 s \gg 1$ (JAIN, 2008). If $\bar{n}_0 s$ is small, the ratchet clicks frequently and a traveling wave approach is more appropriate (ROUZINE *et al.*, 2008). However, a quantitative understanding of the $\bar{n}_0 s \gg 1$ regime is still lacking.

Here, we present a systematic analysis of the problem by first analyzing how selection stabilizes the population against the destabilizing influences of mutation and genetic drift, and later use this insight to derive an approximation to the rate of Muller's ratchet. Before analyzing Eq. (1), it is useful to realize that it implies a common unit of time for the time derivative, the mutation rate and the selection coefficient which is of our choosing (days, months, generations, ...). We can use this freedom to simplify the equation and reveal what the important parameters are that govern the behavior of the equation. In this case, it is useful to use s^{-1} as the unit of time and work with the rescaled time $\tau = ts$. Furthermore, we will formulate the problem in terms of frequencies $x_k = N^{-1}n_k$ rather than numbers of individuals, and obtain

$$\frac{d}{d\tau} x_k = (\bar{k} - k - \lambda)x_k + \lambda x_{k-1} + \sqrt{\frac{x_k}{Ns}} \eta_k \quad (3)$$

where $\lambda = u/s$ is the dimensionless ratio of mutation rate and selection strength. In other words, λ is the average number of mutations that happen over a time s^{-1} (our unit of time). Note that λ uniquely specifies the deterministic part of this equation and its steady state solution $\bar{x}_k = \lambda^k e^{-\lambda}/k!$. The stochastic forces are proportional to $1/\sqrt{Ns}$. Again, the parameter combination $(Ns)^{-1}$ has a simple interpretation as the variance of the stochastic effects accumulated over time s^{-1} . Other than through a prefactor determining the unit of time, any quantity governed by Eq. (3) can depend only on λ and Ns . Hence it is immediately obvious that the ratchet rate cannot depend on $\bar{n}_0 = Ne^{-\lambda}$ alone, but has to depend on $\bar{n}_0 s$ instead (JAIN, 2008). All times and rates in rescaled time units are denoted by greek letters, while we use roman letters for times and rates in units of generations.

Before turning to the ratchet rate, we shall analyze in greater detail the interplay of deterministic and stochastic forces in Eq. (3). A full time dependent analytic solution of the deterministic model was found in (ETHERIDGE *et al.*, 2007). Below, we will present an analytic characterization of the stochastic properties of the system in a limit where stochastic perturbations are small.

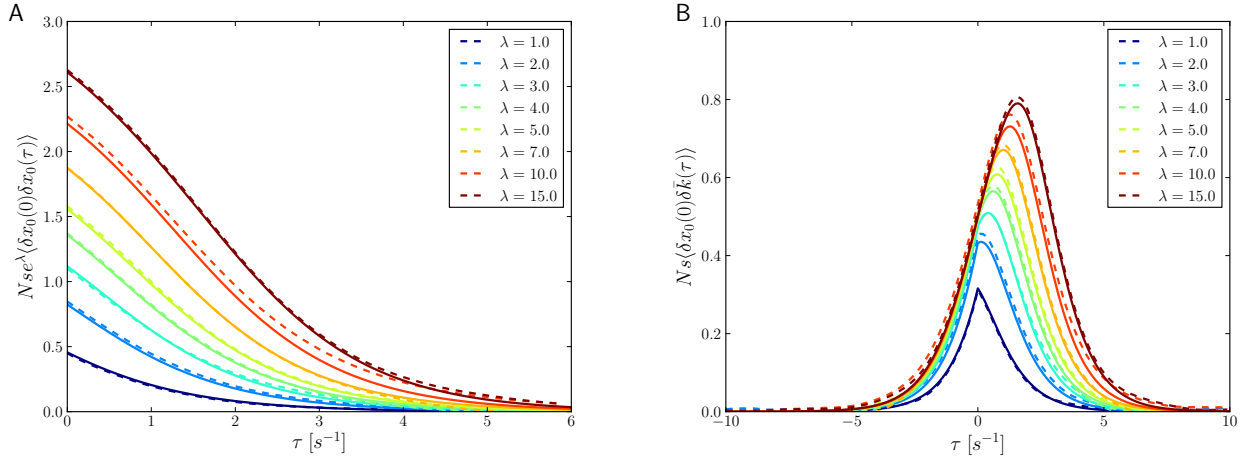


FIG. 2 Panel A: the covariance of the size of the fittest class $x_0(0)$ with $x_0(\tau)$ a time τ later. The normalized auto-correlation of x_0 increases with λ . Panel B: the covariance of $x_0(0)$ with the mean fitness at time τ in the past or future. One observes a pronounced asymmetry, showing that fluctuations of the fittest class propagate towards the bulk of the fitness distribution and results in delayed fluctuations. Simulation results are shown as dashed lines, theory curves are solid. In all cases, $s = 0.01$ and $Nx_0 = 100$. Note that time is measured in units of $1/s$, which is the natural time scale of the dynamics.

Linear stability analysis.

In the limit of large populations, the fluctuations of x_k around the deterministic steady state \bar{x}_k can be analyzed in linear perturbation theory. In other words, we express deviations from the steady state as $\delta x_k = x_k - \bar{x}_k$ and expand

the deterministic part of Eq. (3) to order δx_k^2 . This expansion

$$\frac{d}{d\tau}\delta x_k = -k\delta x_k + \lambda\delta x_{k-1} + \bar{x}_k \sum_{m=0} (m - \lambda)\delta x_m = \sum_m L_{km}\delta x_m \quad (4)$$

defines a linear operator L_{km} . A quick calculation shows that L_{km} has eigenvalues $\kappa_i = -i$ with $i = 0, 1, 2, \dots$. The right eigenvector corresponding to $\kappa_0 = 0$ is simply $\psi_k^{(0)} = \bar{x}_k$, while the right eigenvectors for $i > 0$ are given by

$$\psi_k^{(i)} = \bar{x}_{k-i} - \bar{x}_k \quad (5)$$

where k numbers the coordinate of the vector. This is readily verified by direct substitution (Note that $\bar{x}_i = 0$ for $i < 0$).

The eigenvector $\psi_k^{(0)}$ corresponds to population size fluctuations which in our implementation are controlled by a carrying capacity. The eigenvalue associated with this mode in the computer simulation is large and negative and need not be considered here, see Methods. All other eigenvalues are negative, which is to say that \bar{x}_k is a stable solution.

The eigenvectors for $i > 0$ have an intuitive interpretation: Eigenvector i corresponds to a shift of a fraction of the population by i fitness classes downward. Since such a shift reduces mean fitness, the fittest classes start growing, and undo the shift. More generally, any small perturbation of the population distribution can be expanded into eigenvectors $\delta x_k(\tau) = \sum_j \psi_k^{(j)} a_j(\tau)$ and the associated amplitudes $a_j(\tau)$ will decay exponentially in time with rate j (remember that the unit of time is s^{-1}). Since the amplitudes are projections of δx_k onto the left eigenvectors of L_{mk} , we need to know those as well. For $\kappa_0 = 0$, the left eigenvector is simply $\phi_k^{(0)} = 1$, while the other left eigenvectors are given by

$$\phi_k^{(i)} = \frac{(-1)^{k-i} e^{\lambda} \lambda^{i-k}}{(i-k)!} \quad 0 \leq k \leq i, \quad (6)$$

and $\phi_k^{(i)} = 0$ for $k > i$. With the left and right eigenvectors and the eigenvalue spectrum of the deterministic system on hand, we will now re-instantiate the stochastic part of the dynamics.

$$\frac{d}{d\tau}\delta x_k = \sum_m L_{km}\delta x_m + \sqrt{\frac{\bar{x}_k}{N_s}} \eta_k \quad (7)$$

Note that we approximated the strength of noise by its value at equilibrium. This approximation is justified as long as we consider only small deviations from the equilibrium. The full x_k dependent noise term will be reintroduced later when we turn to Muller's ratchet. Substituting the representation of $\delta x_k(\tau) = \sum_i \psi_k^{(i)} a_i(\tau)$ and projecting onto the left eigenvector $\phi_k^{(j)}$, we obtain the stochastic equations for the amplitudes

$$\frac{d}{d\tau} a_j(\tau) = -j a_j(\tau) + \sum_k \phi_k^{(j)} \sqrt{\frac{\bar{x}_k}{N_s}} \eta_k(\tau). \quad (8)$$

Each noise term η_k contributes to every a_j and induces correlations between the a_j , but each amplitude can be integrated explicitly

$$a_j(\tau) = \int_{-\infty}^{\tau} d\tau' e^{-j(\tau-\tau')} \sum_k \phi_k^{(j)} \sqrt{\frac{\bar{x}_k}{N_s}} \eta_k(\tau'). \quad (9)$$

The covariances of different amplitudes are evaluated in the Supplementary Information and found to be

$$\langle a_i(\tau) a_j(\tau + \Delta\tau) \rangle = \frac{e^{-j\Delta\tau}}{i+j} \sum_k \frac{\phi_k^{(i)} \phi_k^{(j)} \bar{x}_k}{N_s} \quad (10)$$

However, we are not primarily interested in the covariance properties of the amplitudes of eigenvectors, but expect that the fluctuations of the fittest class and fluctuations of the mean fitness are important for the rate of Muller's ratchet and other properties of the dynamics of the population. To this end we express $\delta x_0(\tau)$ and $\bar{k}(\tau)$ as

$$\delta x_0(\tau) = \sum_{j>0} \psi_0^{(j)} a_j(\tau) = -e^{-\lambda} \sum_{j>0} a_j(\tau) \quad (11)$$

$$\delta \bar{k}(\tau) = - \sum_{j>0, k} k \psi_k^{(j)} a_j(\tau) = - \sum_{j>0} j a_j(\tau). \quad (12)$$

Together with Eq. (10), we can now calculate the desired quantities. The calculations required to break down the multiple sums to interpretable expressions are lengthy, but straight-forward and detailed in the supplement. Below, we will present and discuss the results obtained in the supplement.

Fluctuations of x_0 and the mean fitness.

For the variance of the fittest class $\langle \delta x_0^2 \rangle$ and more generally its auto-correlation, we find

$$\begin{aligned} \langle \delta x_0(0) \delta x_0(\tau) \rangle &= \frac{e^{-\lambda}}{Ns} \int_0^1 \frac{d\theta}{\theta} G_\lambda(\theta, \tau) \\ G_\lambda(\theta, \tau) &= e^{\lambda\theta^2 e^{-\tau} - \lambda(1+e^{-\tau})\theta} - e^{-\lambda\theta} - e^{-\lambda\theta e^{-\tau}} + 1 \end{aligned} \quad (13)$$

The variance of the fittest class ($\tau = 0$ in the above expression) is therefore $\sigma^2 = \frac{\bar{x}_0}{Ns} \zeta^2(\lambda)$ where $\zeta^2(\lambda) = \int_0^1 \frac{d\theta}{\theta} G_\lambda(\theta, 0)$ is the standardized variance of the top bin, which depends only on λ . For small λ , it simplifies to $\zeta^2(\lambda) \approx \frac{1}{2}\lambda + \mathcal{O}(\lambda^2)$. This limit corresponds to \bar{x}_0 close to 1 with only a small fraction of the population carrying deleterious mutations $x_1 \approx \lambda = u/s$. The opposite limit of large λ corresponds to a broad fitness distribution where the top class represents only a very small fraction of the entire population. In this limit, the leading behavior of the variance σ^2 is $\sim \frac{\bar{x}_0 \log \lambda}{Ns}$. The full auto-correlation function is shown in Fig. 2A for different values of λ and compared to simulation results, which agree within measurement error. In our rescaled units, the correlation functions decay over a time of order 1, corresponding to a time of order $1/s$ in real time. More precisely, the decay time (in scaled units) increases with increasing λ as $\log \lambda$.

In a similar manner, we can calculate the auto-correlation of the mean fitness

$$\begin{aligned} \langle \delta \bar{k}(0) \delta \bar{k}(\tau) \rangle &= \frac{\lambda e^\lambda}{Ns} \int_0^1 d\theta I_\lambda(\theta, \tau) \\ I_\lambda(\theta, \tau) &= e^{-\tau} e^{\lambda\theta^2 e^{-\tau} - \lambda(1+e^{-\tau})\theta} (\theta + \lambda\theta(\theta e^{-\tau} - 1)(\theta - 1)) \end{aligned} \quad (14)$$

which asymptotes to $(4Ns\bar{x}_0)^{-1}$ for large λ at $\tau = 0$. It is hence inversely proportional to the size of the fittest class x_0 . For large λ , x_0 represents only a tiny fraction of the population and fluctuations of the mean can be substantial even for very large N . This emphasizes the importance of fluctuations of the size of the fittest class for properties of the distribution.

If fluctuations of the mean fitness \bar{k} are driven by fluctuations of the fittest class x_0 , we expect a strong correlation between those fluctuations (ETHERIDGE *et al.*, 2007). Furthermore, fluctuations of x_0 should precede fluctuations of the mean. These expectations are confirmed by the analytic result

$$\langle \delta x_0(0) \delta \bar{k}(\tau) \rangle = \frac{\lambda}{Ns} \int_0^1 d\theta H_\lambda(\theta, \tau) \quad (15)$$

where

$$H_\lambda(\theta, \tau) = \begin{cases} (\theta - 1)e^{-\tau + e^{-\tau}\lambda\theta^2 - \lambda(1+e^{-\tau})\theta} + e^{-\tau - e^{-\tau}\lambda\theta} & \tau > 0 \\ (e^\tau\theta - 1)e^{e^\tau\lambda\theta^2 - \lambda(1+e^\tau)\theta} + e^{-\lambda\theta} & \tau < 0 \end{cases} \quad (16)$$

This expression is shown in Fig. 2B for different values of λ . The cross correlation $\langle \delta x_0(0) \delta \bar{k}(\tau) \rangle$ is asymmetric in time: With larger λ , the peak of the correlation function moves slowly (logarithmically) to larger delays. This result is intuitive, since we expect that fluctuations in the fittest class will propagate to less and less fit classes and that the dynamics of the entire distribution is, at least partly, slaved to the dynamics of the top class.

In all of these three cases, the magnitude of the fluctuations is governed by the parameter Ns , while the shape of the correlation function depends on the parameter λ . Only the unit in which time is measured has to be compared to the strength of selection directly.

The rate of Muller's Ratchet.

The ratchet clicks when the size of the fittest class hits 0, and the rate of the ratchet is given by the inverse of the mean time between successive clicks of the ratchet. Depending on the average size \bar{x}_0 of the fittest class, the model

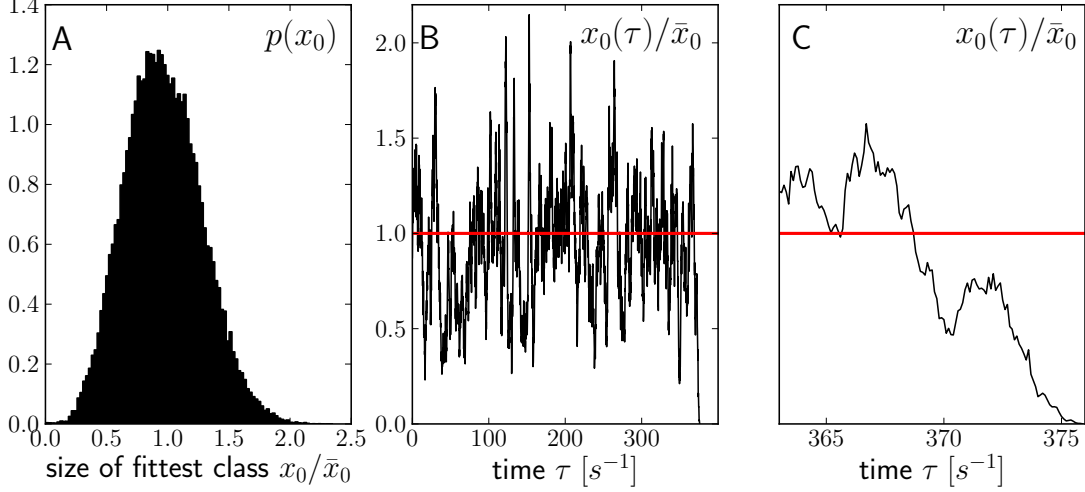


FIG. 3 An example of a click of the ratchet with $N = 5 \times 10^7$, $s = 0.01$ and $\lambda = 10$, corresponding to an average size of the fittest class $\bar{n} \approx 2269$. Panel A: The distribution of x_0 averaged over the time prior to extinction. Panel B&C: The trajectory of $x_0(\tau)$, with the part of the trajectory that ultimately leads to extinction magnified in panel C. The final run towards $x_0 = 0$ takes a few time units, as expected from the results for the correlation functions which suggest a (rescaled) correlation time of $\sim \log \lambda$. Note this time corresponds to a few hundred generations since $s = 0.01$.

displays very different behavior. If $Ns\bar{x}_0$ is comparable to or smaller than 1, the ratchet clicks often without settling to a quasi-equilibrium in between clicks. This limit has been studied in (ROUZINE *et al.*, 2008). Conversely, if $Ns\bar{x}_0 \gg 1$ ratchet clicks are rare and the system stays a long time close to its quasi-equilibrium state \bar{x}_k . Such a scenario, taken from simulations, is illustrated in Fig. 3. Panel A shows the distribution of x_0 prior to the click, while Fig. 3B shows the realized trajectory which ends at $x_0 = 0$. Prior to extinction, $x_0(\tau)$ fluctuates around its equilibrium value and large excursions are rare and short. The final fluctuation which results in the click of the ratchet is zoomed in on in Fig. 3C. Compared to the time the trajectory spends near \bar{x}_0 , the final large excursion away from the steady is short and happens in a few units of rescaled time. Translated back to generations, the final excursion took a few hundred generations ($s = 0.01$ in this example).

In rescaled time, the equation governing the frequency of the top class is

$$\frac{d}{d\tau}x_0(\tau) = \delta\bar{k}(\tau)x_0(\tau) + \sqrt{\frac{x_0(\tau)}{Ns}}\eta_0(\tau) \quad (17)$$

where $\delta\bar{k}(\tau) = \lambda - \bar{k}(\tau)$. The restoring force $\delta\bar{k}$ depends on x_0 , as well as on the size of the other classes x_k . For sufficiently large λ , x_0 is much smaller than x_k with $k \geq 1$, such that the stochastic force is most important for x_0 . The dynamics of x_k , $k \geq 1$, is approximately slaved to the stochastic trajectory of $x_0(\tau)$. We can therefore try to find an approximation of $\delta\bar{k}(\tau)$ in terms of $x_0(\tau)$ only. The linear stability analysis of the mutation selection balance has taught us that the restoring force exerted by the mean fitness on fluctuations in x_0 is delayed with the delay increasing with $\propto \log \lambda$. The latter observation implies that the restoring force on x_0 will mainly depend on the values of x_0 some time of order $\log \lambda$ in the past. Such history dependence complicates the analysis, and this delay has been ignored in previous analysis, which assumed that $\delta\bar{k}(\tau)$ depends on the instantaneous value of $x_0(\tau)$ via $\delta\bar{k}(\tau) = \alpha(1 - x_0(\tau)/\bar{x}_0)$ (GORDO and CHARLESWORTH, 2000; JAIN, 2008; STEPHAN *et al.*, 1993). The parameter α was chosen ad hoc between 0.5 and 0.6. This restoring force is akin to an harmonic potential centered around \bar{x}_0 and the stochastic dynamics is equivalently described by a diffusion equation for the probability distribution $P(x_0, \tau)$.

$$\frac{\partial}{\partial \tau}P(z, \tau) = \frac{1}{2Ns} \frac{\partial^2}{\partial z^2}zP(z, \tau) - \alpha \frac{\partial}{\partial z}(1 - z/\bar{z})zP(z, \tau) \quad (18)$$

where we have denoted x_0 by z for simplicity. The fact that the fittest class is lost whenever its size hits 0 corresponds to an absorbing boundary condition for $P(z, \tau)$ at $z = 0$. For such a one dimensional diffusion problem, the mean first passage time can be computed in closed form (GARDINER, 2004) and this formula has been used in (GORDO and CHARLESWORTH, 2000; STEPHAN *et al.*, 1993) to estimate the rate of the ratchet. An accurate analytic approximation

to that formula has been presented by JAIN (2008). For completeness, we will present an alternative derivation of these results that will help to interpret the more general results presented below. In the limit of interest, $Ns\bar{z} \gg 1$, clicks of the ratchet occur on much longer time scales than the local equilibration of z . We can therefore approximate the distribution as $P(z, \tau) \approx e^{-\gamma\tau}p(z)$ where γ is the rate of the ratchet. In this factorization, $p(z)$ is the quasi-steady distribution shown in Fig. 3A, while γ is the small rate at which $P(z, \tau)$ loses mass due to events like the one shown in Fig. 3C. Inserting this ansatz and integrating Eq. (18) from z to ∞ , we obtain

$$-\gamma P(X > z) = \frac{1}{2Ns} \frac{\partial}{\partial z} z p(z) - \alpha z (1 - z/\bar{z}) p(z) \quad (19)$$

where $P(X > z) = \int_z^\infty dz' p(z')$, which is ≈ 1 for $z < \bar{z}$ and rapidly falls to 0 for $z > \bar{z}$. To obtain the rate γ , we will solve this equation in a regime of small $z \ll \bar{z}$, where the term on the left is important but constant, and in a regime $z \gg (Ns)^{-1}$, where the term on the left can be neglected. For the general discussion below, it will be useful to solve this equation for a general diffusion constant $D(z)$ (here equal to $z/2Ns$), force field $A(z)$ (here equal to $\alpha z(z/\bar{z} - 1)$), and a constant C

$$-C = \frac{\partial}{\partial z} D(z) p(z) + A(z) p(z) \quad (20)$$

with solution

$$p(z) = \frac{1}{D(z)} e^{-\int_0^z dy \frac{A(y)}{D(y)}} \left[\beta - C \int_0^z dy e^{\int_0^y dy' \frac{A(y')}{D(y')}} \right] \quad (21)$$

Note that this solution is inversely proportional to the diffusion constant, while the dependence on selection is accounted for by the exponential factors. For $z \ll \bar{z}$, $C = \gamma$ and $2\alpha Ns \int_0^z dy (1 - z/\bar{z}) \approx 2\alpha Ns z$, such that

$$p(z) \approx \gamma \frac{e^{2Ns\alpha z} - 1}{\alpha z} \quad z \ll \bar{z}, \quad (22)$$

where β is fixed by the boundary condition that $p(z)$ is finite at $z = 0$. Note that $\gamma = p(0)/2Ns$ relates the rate of extinction to $p(0)$. To determine the latter we need to match the $z \ll \bar{z}$ regime to the bulk of the distribution $z \approx \bar{z}$. As can be seen from Eq. (22), the constant term C is unimportant in this regime ($e^{2Ns\alpha z} \gg 1$). Setting $C = 0$ in Eq. (21), we find

$$p(z) \approx \frac{\bar{z}}{z} \frac{\sqrt{Ns\alpha}}{\sqrt{\bar{z}\pi}} e^{-\alpha Ns \frac{(z-\bar{z})^2}{\bar{z}}} \quad zNs\alpha \gg 1. \quad (23)$$

The integration constant β in Eq. (21) is fixed by the normalization. Since $p(z)$ is concentrated around $z = \bar{z}$ and has a Gaussian shape around \bar{z} , the normalization factor is simply $1/\sqrt{2\pi\sigma^2}$, where $\sigma^2 = \bar{z}/2\alpha Ns$ is the variance of the Gaussian. The factor \bar{z}/z corresponds to the $1/D(z)$ term, scaled so that it equals 1 in the vicinity of \bar{z} . Note that we have already calculated the variance of $p(z)$ earlier, Eq. (13), and that consistency with this result would require that α is determined by Eq. (13).

The two approximate solutions Eq. (22) and Eq. (23) are both accurate in the intermediate regime $(Ns\alpha)^{-1} \ll z \ll \bar{z}$, which allows us to determine the rate γ in Eq. (22) by matching the two solutions. This matching implies that

$$\gamma = \frac{\sqrt{\bar{z}Ns\alpha^3}}{\sqrt{\pi}} e^{-\alpha Ns e^{-\lambda}} \quad (24)$$

which agrees with results obtained previously (JAIN, 2008). Note that this rate only depends on the parameters λ and Ns of the rescaled model. Since rates have units of inverse time, this expression has to be multiplied by s to obtain the rate in units of inverse generations.

However, Eq. (24) does not describe the rate accurately, as is obvious from the comparison with simulation results shown in Fig. 4A. The plot shows the rescaled ratchet rate $\gamma \times \sqrt{\pi}/\sqrt{\bar{z}Ns\alpha^3}$, which according to Eq. (24) should be simply $\exp[-\alpha Ns e^{-\lambda}]$, indicated by the black line. The plot shows clearly that the simulation results often differ from the prediction of Eq. (24) by a large factor. It seems as if α needs to depend on λ , as we already noticed above when comparing the variance of $p(z)$ to Eq. (13). In fact, fixing α via Eq. (13) improves the agreement substantially, but still does not describe the simulations quantitatively.

The reason for the discrepancy is the time delay between $\delta\bar{k}$ and z which we quantified by calculating the correlation between $\delta\bar{k}(\tau)$ and $z(\tau + \Delta\tau)$. Hence we cannot use an approximation where $\delta\bar{k}$ depends on the instantaneous value of

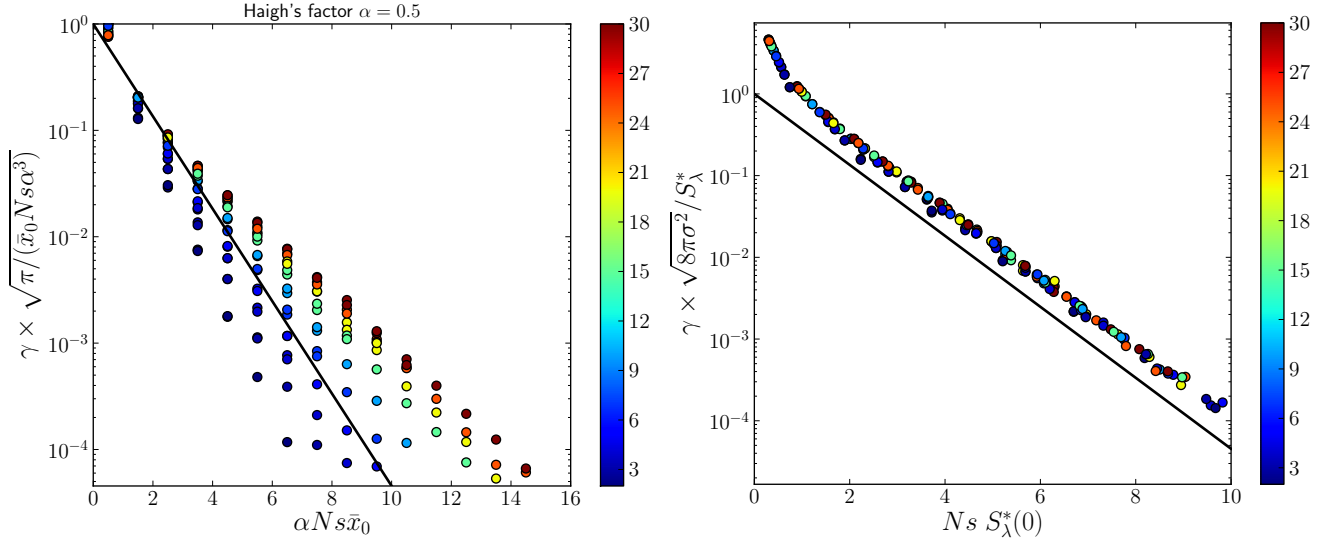


FIG. 4 The ratchet rate from simulation vs prediction. Both panels show the ratchet rate γ , rescaled with a prefactor to isolate the exponential dependence predicted by analytic approximations; λ is color-coded. Panel A compares the simulation results to the prediction of Eq. (24), which is shown as a straight line. The approximation works only for a particular value of λ , for otherwise the exponential dependence on $N\bar{x}_0s$ is not predicted correctly. Panel B compares simulation results to the prediction of Eq. (32), again indicated by the straight line. The exponential dependence of rate on $NsS_\lambda^*(0)$ is well confirmed by simulation results.

z , but have to calculate $\delta\bar{k}$ from the past trajectory of z . If the fittest class is the only one that is strongly stochastic, we can calculate $\delta\bar{k}(\tau)$ for a given trajectory $z(\rho)$, $\rho \leq \tau$ by integrating the deterministic evolution equations for x_k with $k \geq 1$ with $z(\rho)$ as an external forcing.

Eq. (17) now not only depends on $z(\tau)$, but on all $z(\rho)$ with $\rho \leq \tau$ and cannot be mapped to a diffusion equation. Nevertheless, it corresponds to a well defined stochastic integral, known as a path integral in physics (FEYNMAN and HIBBS, 1965), which is amenable to systematic numerical approximation. To introduce path integrals, it is useful to discretize Eq. (17) in time and express $z(\rho_i)$ in terms of the state at time $\rho_{i-1} = \rho_i - \Delta\tau$ and the earlier time points. For simplicity, we will use the notation z_i for $z(\rho_i)$.

$$z_i - z_{i-1} = \Delta\tau \delta\bar{k}_{i-1} z_{i-1} + \sqrt{\frac{z_{i-1} \Delta\tau}{Ns}} \eta_{i-1} \quad (25)$$

where $\delta\bar{k}_{i-1}$ depends on all previous time points ρ_j with $j < i$. In the limit $\Delta\tau \rightarrow 0$, this difference equation converges against Eq. (17) interpreted in the Itô sense since the z dependent prefactor of the noise term is evaluated at ρ_{i-1} rather than at an intermediate time point between ρ_{i-1} and ρ_i . We can express this transition probability $P_\tau(z|z_0)$ between the initial state z_0 and the final state $z = z_m$ as a series of integrals over all intermediate states z_i for $0 < 1 < m$.

$$P_\tau(z|z_0) = \int \prod_{i=1}^{m-1} dz_i P_{\Delta\tau}(z|\{z_j\}_{j<m}) P_{\Delta\tau}(z_{m-1}|\{z_j\}_{j<m-1}) \cdots P_{\Delta\tau}(z_1|z_0) \quad (26)$$

Each of these infinitesimal transitions correspond to solutions of Eq. (25) with η_i drawn from a standard Gaussian (LAU and LUBENSKY, 2007). Hence

$$P_{\Delta\tau}(z_i|\{z_j\}_{j<i}) = \frac{\sqrt{Ns}}{\sqrt{2\pi\Delta\tau z_{i-1}}} \exp \left[-Ns \frac{(z_i - z_{i-1} - \Delta\tau z_{i-1} \delta\bar{k}_{i-1})^2}{2\Delta\tau z_{i-1}} \right] \quad (27)$$

In the limit of many intermediate steps and small $\Delta\tau$, the transition probability can therefore be written as

$$P_\tau(z|z_0) = \int \mathcal{D}z_\rho \exp \left[-Ns \int_0^\tau d\rho \frac{[\dot{z}_\rho - z_\rho \delta\bar{k}_\rho]^2}{2z_\rho} \right] = \int \mathcal{D}z_\rho e^{-Ns S_\lambda(\{z_\rho\})} \quad (28)$$

where $\mathcal{D}z_\rho$ is the limit of $\prod_{i=1}^m dz_i (2\pi\Delta\tau z_{i-1}/Ns)^{-1/2}$ known as the path integral measure, and we have replaced the discrete time index by its continuous analog. The path integral extends over all continuous path connecting the endpoints z_0 and $z_\tau = z$. The functional $S_\lambda(\{z_\rho\})$ in the exponent closely corresponds to the “action” in physics (FEYNMAN and HIBBS, 1965) which is minimized by classical dynamics. Here minimization of the “action” defines the most likely trajectory. Note that $S_\lambda(\{z_\rho\})$ depends on the entire path $\{z_\rho\}$ with $0 \leq \rho \leq \tau$, while the functional itself only depends on λ . The strength of genetic drift appears as a prefactor of $S_\lambda(\{z_\rho\})$ in the exponent.

The most likely path z_ρ^* connecting the end-points points z_0 and z_τ in time τ can be determined either by solving the Euler-Lagrange equations or by numerical minimization, see below. Along with the functional, z_ρ^* depends only on λ . Given this extremal path, we can parameterize every other path connecting z_0 and z_τ as $z_\rho = z_\rho^* + \delta z_\rho$, where δz_ρ vanishes at both endpoints ($\delta z_0 = \delta z_\tau = 0$). Denoting the minimal action associated with z_ρ^* by $S_\lambda^*(z_\tau, z_0)$, we have

$$P_\tau(z_\tau|z_0) = e^{-NsS_\lambda^*(z_\tau, z_0)} \int \mathcal{D}\delta z_\rho e^{-Ns \delta S_\lambda(\{\delta z_\rho\}, z_\tau, z_0)} = \mathcal{N}^{-1} e^{-NsS_\lambda^*(z_\tau, z_0)} \quad (29)$$

where \mathcal{N}^{-1} factor is equal to the integral over the fluctuations, which in general depends on z_ρ^* . The prefactor Ns in $e^{-Ns \delta S_\lambda(\{\delta z_\rho\}, z_\tau, z_0)}$ implies that deviations from the optimal path are suppressed in large populations. If $\delta S_\lambda(\{\delta z_\rho\}, z_\tau, z_0)$ is independent of the final point z_τ , \mathcal{N} can be determined by the normalizing $P_\tau(z_\tau|z_0)$ with respect to z_τ . In the general case, calculating the fluctuation integral is difficult, and we will determine it here by analogy to the history independent solution presented above Eq. (18)–(24).

If the stochastic dynamics admits an (approximately) stationary distribution, $P_\tau(z_\tau|z_0)$ becomes independent of τ and z_0 and coincides with the steady state probability distribution $p(z)$. It therefore becomes the analog of Eq. (23), which for arbitrary diffusion equations is given by the inverse diffusion constant (the prefactor z^{-1}), multiplied by an exponential quantifying the trade-off between deterministic and stochastic forces. In this path integral representation, the exponential part is played by $e^{-NsS_\lambda^*(z)}$, where $S_\lambda^*(z)$ is a function of the final point z and λ only. The prefactor is independent of the selection term and can hence be determined through the analogy to the Markovian case discussed above

$$p(z) \approx \frac{\bar{z}}{z} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-NsS_\lambda^*(z)} \quad Ns z \gg 1 \quad (30)$$

where the normalization is obtained by assuming an approximately Gaussian distribution around the steady state value \bar{z} and the variance σ^2 is given by Eq. (13). Note that this solution is not valid very close to the absorbing boundary since this boundary is not accounted for by the path integral, at least not without some special care. As in the history independent case discussed above, this approximate distribution should be thought of as the time independent “bulk” distribution in $P(z, \tau) = e^{-\gamma\tau} p(z)$. To determine the rate extinction rate γ , we again need to understand how probable it is that a trajectory actually hits $z = 0$, given that it has come pretty close.

To this end, we need a local solution of Eq. (17) in the boundary layer $z \ll \bar{z}$ as already obtained for the history independent case in Eq. (22). Once a trajectory comes close to $z = 0$, its fate, i.e., whether it goes extinct or returns to $z \approx \bar{z}$, is decided quickly. Hence we can make an instantaneous approximation for $\delta\bar{k}$ which does depend on the past trajectory, but for the time window under consideration it is simply a constant, α , yet to be determined. Having reduced the problem to Eq. (22) we can determine α , and hence γ , by matching of the boundary solution to Eq. (30) in the regime $(Ns)^{-1} \ll z \ll \bar{z}$ where both are accurate. The matching condition is

$$\frac{\bar{z}}{z} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-Ns(S_\lambda^*(0) + z\partial_z S_\lambda^*(z)|_{z=0})} = \frac{\gamma}{\alpha z} e^{2\alpha Ns z} \quad (31)$$

which determines α and γ by the matching requirement $2\alpha = -\partial_z S_\lambda^*(z)|_{z=0}$ and

$$\gamma = \frac{|\partial_z S_\lambda^*(z)|_{z=0} \bar{z} e^{-NsS_\lambda^*(0)}}{2\sqrt{2\pi\sigma^2}} \approx \frac{S_\lambda^*(0)}{\sqrt{8\pi\sigma^2}} e^{-NsS_\lambda^*(0)} \quad (32)$$

where we approximated $|\partial_z S_\lambda^*(z)|_{z=0} \approx S_\lambda^*(0)/\bar{z}$. The variance σ^2 is given by Eq. (13) and depends on λ and Ns as $\sigma^2 = \bar{z}\zeta^2(\lambda)/Ns$. Note that γ is in units of s and needs to be multiplied by s for conversion to units of inverse generations. In contrast to Markovian case above, the variance of the “bulk” is no longer simply related to strength of selection near the $z = 0$ “boundary”.

Since we don’t know how to calculate $S_\lambda^*(0)$ or the most likely path z_ρ^* analytically, we determined discrete approximations to z_ρ^* numerically as described in section Model and Methods. Examples of numerically determined most likely path and the corresponding trajectory of the mean fitness are shown in Fig. 5 for different values of λ . Generically, we find a rapid reduction of the size z of the size of the fittest class such that the mean fitness has only

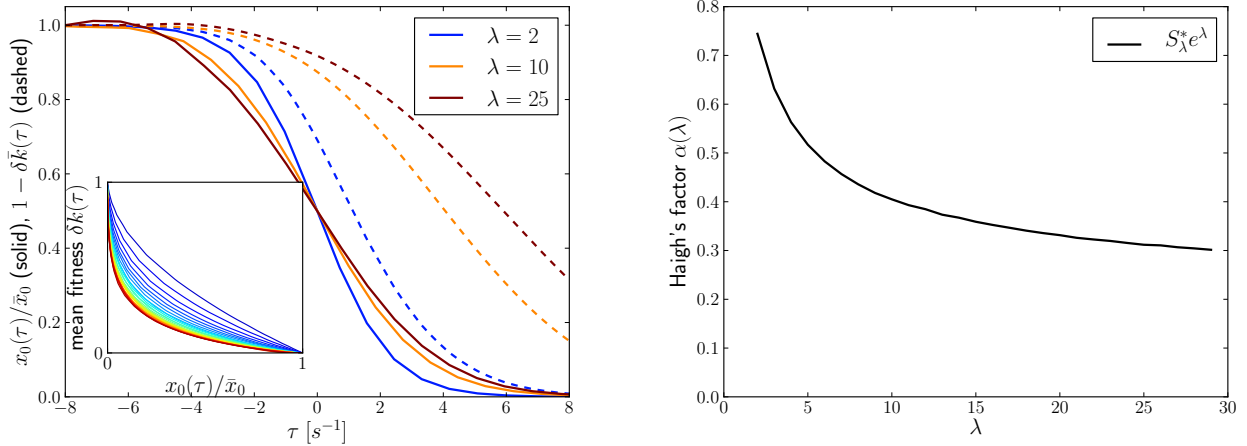


FIG. 5 Panel A: The most likely path $x_0^*(\tau)$ to extinction of the fittest class and the concomitant reduction of the mean fitness for different λ are plotted against time. Times are shifted such that $x_0^*(0) = \frac{x_0}{2}$. The inset shows the mean fitness $\delta \bar{k}(\tau)$ plotted against $x_0(\tau)/\bar{x}_0$ for different values of λ . Panel B: “Haigh’s factor” $\alpha(\lambda) = S_\lambda^*(0)e^\lambda$ as a function of λ determined numerically.

partially responded. The inset of Fig. 5A shows how changes in mean fitness $\delta \bar{k}$ are related to z for different λ . For large λ , the mean fitness changes only very slowly with z , which increases the probability of large excursions and hence the rate of the ratchet.

This numerically determined minimal action $S_\lambda^*(0)$ together with the approximation Eq. (32) describes the rate of the ratchet, as determined in simulations, extremely well. Fig. 4B shows the same simulation data as Fig. 4A, but this time rescaled by $\sqrt{8\pi\sigma^2(\lambda)}/S_\lambda^*(0)$ as a function of $NsS_\lambda^*(0)$. After this rescaling, we expect all data points to lie on the same curve given by $\exp[-NsS_\lambda^*(0)]$, as is indeed found for many different values of u , s and N with $\lambda = u/s$ ranging from 1 to 30. Note that the vertical shift of the black line relative to the data points depends on the prefactor, which we have approximated. Hence we should not expect agreement better than to about a factor of 2. The important point is that the exponential dependence of the rate on $NsS_\lambda^*(0)$ is correctly captured by Eq. (32).

Previous studies of Muller’s ratchet suggested that the rate depends exponentially on $\alpha N s e^{-\lambda}$ (JAIN, 2008). To relate this to our results, we determined “Haigh’s factor” α numerically from $S_\lambda^*(0)$ and plotted it in Fig. 5B. We find that $\alpha(\lambda)$ drops from around 0.8 to 0.3 as λ increases from 1 to 30. The previously used values 0.5 – 0.6 for α correspond to $\lambda \approx 6$. Using $\alpha(\lambda) = S_\lambda^*(0)e^\lambda$ as shown in Fig. 5B, we can recast Eq. (32) into its traditional form and undo the scaling with s . In units of generations, the mean time between clicks is given by

$$T_{\text{click}} \approx \frac{2.5\zeta(\lambda)}{\alpha(\lambda)s\sqrt{Nse^{-\lambda}}} e^{Ns\alpha(\lambda)e^{-\lambda}} \quad (33)$$

where $\zeta(\lambda)$ is determined by Eq. (13) and the factor 2.5 is introduced to approximate the part of the prefactor that is independent of N , s or λ . A direct comparison of this expression with simulation results is shown in supplementary figure 1.

III. CONCLUSION.

The main difficulty impeding better understanding of even simple models of evolution is the fact that rare events involving a few or even single individuals determine the fate of the entire population. The important individuals are those in the high fitness tail of the distribution. Fluctuations in the high fitness tail propagate towards more mediocre individuals which dominate a typical population sample.

We have analyzed the magnitude, decay, and propagation of fluctuations of the fitness distribution in a simple model of the balance between deleterious mutations and selection. In this model, individuals in the fittest class evolve approximately neutrally. Fluctuations in the size of this class propagate to the mean, which in turn generates a delayed restoring force opposing the fluctuation. We have shown that the variance of the fluctuations in the population n_0 of the top bin is proportional to n_0/s and increases as $\log \lambda$ with the ratio λ of the mutation rate u and the mutational effect s . Fluctuations of n_0 perturb the mean after a time $\sim s^{-1} \log \lambda$. These two observations have a straightforward

connection: sampling fluctuations can accumulate without a restoring force for a time $s^{-1} \log \lambda$. During this time, the typical perturbation of the top bin by drift is $\sim \sqrt{n_0 s^{-1} \log \lambda}$ and hence the variance is $\approx n_0 s^{-1} \log \lambda$. We have used these insights into the coupling between n_0 , the mean fitness, and the resulting delayed restoring force on fluctuations of n_0 to approximate the rate of Muller's ratchet.

The history dependence of the restoring force has not been accounted for in previous analysis of the rate of Muller's ratchet (GORDO and CHARLESWORTH, 2000; HAIGH, 1978; JAIN, 2008; STEPHAN *et al.*, 1993) who introduced a constant factor to parameterize the effective strength of the selection opposing fluctuations in the top bin, or WAXMAN and LOEWE (2010), who replaced all mutant classes by one effective class and thereby mapped the problem to the fixation of a deleterious allele. We have shown that to achieve agreement between theory and numerical simulation one must account for the delayed nature of selection acting on fluctuations. Comparing our final expression for the ratchet rate with that given previously (GORDO and CHARLESWORTH, 2000; JAIN, 2008; STEPHAN *et al.*, 1993), the history dependence manifests itself as a decreasing effective strength of selection with increasing $\lambda = u/s$. This decrease is due to a larger temporal delay of the response of the mean fitness to fluctuations of the least loaded class. History dependence is a general consequence of projecting a multi-dimensional stochastic dynamics onto a lower dimensional space (here, the size n_0 of the fittest class). Such memory effects can be accounted for by the path-integral formulation of stochastic processes which we used to approximate the rate of Muller's ratchet.

Even though the model is extremely simplistic and the sensitive dependence of the ratchet rate on poorly known parameters such as the effect size of mutations, population size, and mutation rate, precludes quantitative comparison with the real world, we believe that some general lessons can be learned from our analysis. The propagation of fluctuations from the fittest to less fit individuals is expected to be a generic feature of many models and natural populations. In particular, very similar phenomena arise in the dynamics of adapting populations driven by the accumulation of beneficial mutations (COHEN *et al.*, 2005; DESAI and FISHER, 2007; HALLATSCHEK, 2011; NEHER *et al.*, 2010; ROUZINE *et al.*, 2008, 2003; TSIMRING *et al.*, 1996). The speed of these traveling waves is typically determined by stochastic effects at the high fitness edges. We expect that the fluctuations of the speed of adaptation can be understood and quantified with the concepts and tools that we introduced above.

Populations spread out in fitness have rather different coalescence properties than neutral populations, which are described by Kingman's coalescent (KINGMAN, 1982). These differences go beyond the familiar reduction in effective population size and distortions of genealogies due to background selection (CHARLESWORTH *et al.*, 1993; HIGGS and WOODCOCK, 1995; WALCZAK *et al.*, 2011). The most recent common ancestor of such populations most likely derives from this high fitness tail and fluctuations of this tail determine the rate at which lineages merge and thereby the genetic diversity of the population (BRUNET *et al.*, 2007; NEHER and SHRAIMAN, 2011; ROUZINE and COFFIN, 2007). Thus, quantitative understanding of fluctuations of fitness distributions is also essential for understanding non-neutral coalescent processes.

Generalizing the analysis of fluctuations of fitness distributions to adapting "traveling waves" and the study of their implications for the coalescent properties of the population are interesting avenues for future research.

IV. ACKNOWLEDGEMENTS

We are grateful for stimulating discussions with Michael Desai, Dan Balick and Sid Goyal. RAN is supported by an ERC-starting grant HIVEVO 260686 and BIS acknowledges support from NIH under grant GM086793. This research was also supported in part by the NSF under Grant No. NSF PHY11-25915.

References

- BRUNET, E., B. DERRIDA, A. H. MUELLER, and S. MUNIER, 2007 Effect of selection on ancestry: an exactly soluble case and its phenomenological generalization. *Physical review E, Statistical, nonlinear, and soft matter physics* **76**: 041104.
- CHARLESWORTH, B., 2012 The effects of deleterious mutations on evolution at linked sites. *Genetics* **190**: 5–22.
- CHARLESWORTH, B., M. T. MORGAN, and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–303.
- COHEN, E., D. A. KESSLER, and H. LEVINE, 2005 Front propagation up a reaction rate gradient. *Phys Rev E Stat Nonlin Soft Matter Phys* **72**: 066126.
- DESAI, M. M. and D. S. FISHER, 2007 Beneficial mutation selection balance and the effect of linkage on positive selection. *Genetics* **176**: 1759–98.
- ETHERIDGE, A., P. PFAFFELHUBER, and A. WAKOLBINGER, 2007 How often does the ratchet click? facts, heuristics, asymptotics. In *Trends in Stochastic Analysis*, edited by P. M. Jochen Blath and M. Scheutzow, pp. 365–390, Cambridge University Press 2009.
- FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. *Genetics* **78**: 737–56.
- FEYNMAN, R. P. and A. R. HIBBS, 1965 *Quantum mechanics and Path Integrals*. McGraw-Hill Inc, New York.

- GARDINER, C. W., 2004 *Handbook of stochastic methods for Physics, Chemistry and the Natural sciences*. Springer.
- GESSLER, D. D., 1995 The constraints of finite size in asexual populations and the rate of the ratchet. *Genet Res* **66**: 241–53.
- GORDO, I. and B. CHARLESWORTH, 2000 The degeneration of asexual haploid populations and the speed of Muller’s ratchet. *Genetics* **154**: 1379–87.
- GOYAL, S., D. J. BALICK, E. R. JERISON, R. A. NEHER, B. I. SHRAIMAN, and M. M. DESAI, 2012 Dynamic mutation selection balance as an evolutionary attractor. *Genetics* .
- GRADSHTEYN, I. S. and I. M. RYZHIK, 2007 *Table of Integrals, Series, and Products*. Academic Press, New York.
- HAIGH, J., 1978 The accumulation of deleterious genes in a population – Muller’s ratchet. *Theoretical Population Biology* **14**: 251–67.
- HALLATSCHEK, O., 2011 The noisy edge of traveling waves. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 1783–7.
- HIGGS, P. and G. WOODCOCK, 1995 The accumulation of mutations in asexual populations and the structure of genealogical trees in the presence of selection. *J. Math. Biol.* **33**: 677–102.
- JAIN, K., 2008 Loss of least-loaded class in asexual populations due to drift and epistasis. *Genetics* **179**: 2125–34.
- KINGMAN, J., 1982 On the genealogy of large populations. *Journal of Applied Probability* **19** IS -: 27–43.
- LAU, A. W. C. and T. C. LUBENSKY, 2007 State-dependent diffusion: Thermodynamic consistency and its path integral formulation. *Phys Rev E Stat Nonlin Soft Matter Phys* **76**: 011123.
- LYNCH, M., R. BÜRGER, D. BUTCHER, and W. GABRIEL, 1993 The mutational meltdown in asexual populations. *J Hered* **84**: 339–44.
- MULLER, H. J., 1964 The relation of recombination to mutational advance. *Mutat Res* **106**: 2–9.
- NEHER, R. A. and B. I. SHRAIMAN, 2011 Genetic draft and quasi-neutrality in large facultatively sexual populations. *Genetics* **188**: 975–996.
- NEHER, R. A., B. I. SHRAIMAN, and D. S. FISHER, 2010 Rate of adaptation in large sexual populations. *Genetics* **184**: 467–481.
- OLIPHANT, T., 2007 Python for scientific computing. *Computing in Science & Engineering* **9**: 10–20.
- PFÄFFELHUBER, P., P. R. STAAB, and A. WAKOLBINGER, 2012 Muller’s ratchet with compensatory mutations. to appear in *Annals of Applied Probability* xxx, 26 pages, 3 figures.
- RICE, W. R., 1987 Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* **116**: 161–7.
- ROUZINE, I. M., E. BRUNET, and C. O. WILKE, 2008 The traveling-wave approach to asexual evolution: Muller’s ratchet and speed of adaptation. *Theoretical Population Biology* **73**: 24–46.
- ROUZINE, I. M. and J. M. COFFIN, 2007 Highly fit ancestors of a partly sexual haploid population. *Theoretical Population Biology* **71**: 239–50.
- ROUZINE, I. M., J. WAKELEY, and J. M. COFFIN, 2003 The solitary wave of asexual evolution. *Proc Natl Acad Sci USA* **100**: 587–92.
- STEPHAN, W., L. CHAO, and J. G. SMALE, 1993 The advance of Muller’s ratchet in a haploid asexual population: approximate solutions based on diffusion theory. *Genet Res* **61**: 225–31.
- STEPHAN, W. and Y. KIM, 2002 Recent applications of diffusion theory to population genetics. In *Modern Developments in Theoretical Population Genetics*, edited by M. Slatkin and M. Veuille, pp. 72–93, Oxford University Press, Oxford, UK.
- TSIMRING, L., H. LEVINE, and D. KESSLER, 1996 RNA virus evolution via a fitness-space model. *Phys Rev Lett* **76**: 4440–4443.
- WALCZAK, A. M., L. E. NICOLAISEN, J. B. PLOTKIN, and M. M. DESAI, 2011 The structure of genealogies in the presence of purifying selection: A ”fitness-class coalescent”. *Genetics* .
- WAXMAN, D. and L. LOEWE, 2010 A stochastic model for a single click of muller’s ratchet. *Journal of Theoretical Biology* **264**: 1120–1132.

Appendix A: Supplement

To analyze the covariances of different fitness classes, we begin with Eq. (7) of the main text, which expresses $\delta x_k(\tau)$ in terms of eigenvectors $\delta x_k(\tau) = \sum_j \psi_k^{(j)} a_j(\tau)$. Projecting on the left eigenvectors then results in equations for $a_j(\tau)$

$$\frac{d}{d\tau} a_j(\tau) = -j a_j(\tau) + \sum_k \phi_k^{(j)} \sqrt{\frac{\bar{x}_k}{Ns}} \eta_k(\tau) \quad (\text{A1})$$

where $\eta_k(\tau)$ are uncorrelated Gaussian white noise terms with $\langle \eta_k(\tau) \eta_l(\tau') \rangle = \delta_{kl} \delta(\tau - \tau')$. Since each noise term η_k contributes to all a_j , the noise induces correlated fluctuations of the $a_j(\tau)$, which we need to understand in order to analyze the fluctuations of the fitness distributions. The inhomogeneous Eq. (A1) has the solution

$$a_j(\tau) = \int_{-\infty}^{\tau} d\tau' e^{-j(\tau-\tau')} \sum_k \phi_k^{(j)} \sqrt{\frac{\bar{x}_k}{Ns}} \eta_k(\tau') \quad (\text{A2})$$

The autocorrelation function of the loadings of different eigendirections separated by $\Delta\tau$ in time is therefore given by

$$\begin{aligned} \langle a_i(\tau) a_j(\tau + \Delta\tau) \rangle &= \int_{-\infty}^{\tau} d\tau' \int_{-\infty}^{\tau+\Delta\tau} d\tau'' e^{-i(\tau-\tau')-j(\tau+\Delta\tau-\tau'')} \sum_{k,l} \frac{\phi_k^{(i)} \phi_l^{(j)} \sqrt{\bar{x}_k \bar{x}_l}}{Ns} \langle \eta_k(\tau') \eta_l(\tau'') \rangle \\ &= \int_{-\infty}^{\tau} d\tau' e^{-i(\tau-\tau')-j(\tau+\Delta\tau-\tau')} \sum_k \frac{\phi_k^{(i)} \phi_k^{(j)} \bar{x}_k}{Ns} \\ &= \frac{e^{-j\Delta\tau}}{i+j} \sum_k \frac{\phi_k^{(i)} \phi_k^{(j)} \bar{x}_k}{Ns} \end{aligned} \quad (\text{A3})$$

where we have used $\langle \eta_k(\tau) \eta_l(\tau') \rangle = \delta_{kl} \delta(\tau - \tau')$.

Correlation functions n_0 and the mean

To calculate the variances and covariance of x_0 and the mean fitness, we express them in terms of the eigenmodes $a_j(\tau)$ ($j > 0$)

$$\delta x_0(\tau) = \sum_{j>0} \psi_0^{(j)} a_j(\tau) = -e^{-\lambda} \sum_{j>0} a_j(\tau) \quad (\text{A4})$$

$$\delta \bar{k}(\tau) = - \sum_{j>0,k} k \psi_k^{(j)} a_j(\tau) = - \sum_{j>0,k} k (\bar{x}_{k-j} - \bar{x}_k) a_j(\tau) = - \sum_{j>0} j a_j(\tau) \quad (\text{A5})$$

The auto-correlation of x_0

The auto-correlation of x_0 is given by

$$\begin{aligned} \langle x_0(\tau) x_0(\tau + \Delta\tau) \rangle &= e^{-2\lambda} \sum_{i,j>0} \frac{e^{-j\Delta\tau}}{i+j} \sum_k \frac{\phi_k^{(i)} \phi_k^{(j)} \bar{x}_k}{Ns} \\ &= \frac{e^{-\lambda}}{Ns} \sum_{i,j>0} \frac{\lambda^{i+j} e^{-j\Delta\tau}}{i+j} \sum_{k=0}^{\min(i,j)} \frac{(-1)^{i+j} \lambda^{-k}}{(j-k)!(i-k)!k!} \\ &= \frac{e^{-\lambda}}{Ns} \int_0^{\infty} dz \sum_{i,j>0} e^{-z(i+j)} \lambda^{i+j} e^{-j\Delta\tau} \sum_{k=0}^{\min(i,j)} \frac{(-1)^{i+j} \lambda^{-k}}{(j-k)!(i-k)!k!} \end{aligned} \quad (\text{A6})$$

Let us focus on the triple sum inside the integral and simplify it by introducing $a = -\lambda e^{-z}$ and $b = -\lambda e^{-z-\Delta\tau}$. Furthermore, let us look at the $i = j$ and the $i \neq j$ contributions separately. The diagonal contribution ($i = j$) is

$$\begin{aligned}
& \sum_{i=0} a^i b^i \sum_{k=0}^i \frac{\lambda^{-k}}{(i-k)!(i-k)!k!} = \sum_{i>0} \sum_{k=0}^i \frac{(ab)^{i-k} (ab)^k \lambda^{-k}}{k!((i-k)!)^2} \\
& = \sum_{k>0} \sum_{i \geq k} \frac{(ab)^{i-k} (ab)^k \lambda^{-k}}{k!((i-k)!)^2} + \sum_{i>0} \frac{(ab)^i}{(i!)^2} \\
& = \sum_{k>0} \frac{(ab)^k \lambda^{-k}}{k!} \sum_{n \geq 0} \frac{(ab)^n}{(n!)^2} + J_0(-\iota 2\sqrt{ab}) - 1 \quad \text{using } n = i - k \\
& = (e^{ab/\lambda} - 1) J_0(-\iota 2\sqrt{ab}) + J_0(-\iota 2\sqrt{ab}) - 1 \\
& = e^{ab/\lambda} J_0(-\iota 2\sqrt{ab}) - 1
\end{aligned} \tag{A7}$$

where $J_n(z)$ is the n th Bessel function of first kind, and $\iota = \sqrt{-1}$. When evaluating the off-diagonal contribution, we will encounter terms like

$$\sum_{k>0} \frac{(ab)^k}{k!(k+m)!} = \sum_k \frac{(ab)^k}{k!(k+m)!} - \frac{1}{m!} = \frac{J_m(2\iota\sqrt{ab})}{(\iota\sqrt{ab})^m} - \frac{1}{m!} \tag{A8}$$

The off-diagonal contribution can be further split into the parts $i > j$ and $i < j$ which can be evaluated as follows:

$$\begin{aligned}
& \sum_{0 < i < j} a^i b^j \sum_{k \geq 0} \frac{\lambda^{-k}}{k!(j-k)!(i-k)!} = \sum_{i>0} \sum_{j>i} a^i b^{j-i+i} \sum_{k \geq 0} \frac{\lambda^{-k}}{k!(i+(j-i)-k)!(i-k)!} \\
& = \sum_{i>0} \sum_{m>0} a^i b^{m+i} \sum_{k \geq 0} \frac{\lambda^{-k}}{k!(i+m-k)!(i-k)!} \quad \text{using } m = j - i \\
& = \sum_{k>0} \sum_{m>0} \sum_{i \geq k} \frac{a^i b^{m+i} \lambda^{-k}}{k!(i+m-k)!(i-k)!} + \sum_{m>0} \sum_{i>0} \frac{a^i b^{m+i}}{(i+m)!i!} \\
& = \sum_{k>0} \sum_{m>0} \sum_{n \geq 0} \frac{a^{n+k} b^{m+n+k} \lambda^{-k}}{k!(n+m)!n!} + \sum_{m>0} b^m \sum_{i>0} \frac{(ab)^i}{(i+m)!i!} \quad \text{using } n = i - k \\
& = \sum_{k>0} \sum_{m>0} \frac{a^k b^{m+k} \lambda^{-k}}{k!} \frac{J_m(2\iota\sqrt{ab})}{(\iota\sqrt{ab})^m} + \sum_{m>0} b^m \left(\frac{J_m(2\iota\sqrt{ab})}{(\iota\sqrt{ab})^m} - \frac{1}{m!} \right) \\
& = \sum_{m>0} \sum_{k \geq 0} \frac{a^k b^k b^{m/2} a^{-m/2} \lambda^{-k}}{k!} \frac{J_m(2\iota\sqrt{ab})}{(\iota)^m} - e^b + 1 = \sum_{m>0} \left(\frac{b}{a} \right)^{m/2} e^{ab/\lambda} \frac{J_m(2\iota\sqrt{ab})}{(-1)^{m/2}} - e^b + 1 \\
& = e^{ab/\lambda} \sum_{m>0} \left(\frac{b}{a} \right)^{m/2} (-\iota)^m J_m(2\iota\sqrt{ab}) - e^b + 1
\end{aligned} \tag{A9}$$

The off-diagonal terms for $i > j$ is obtained by interchanging a and b such that the full off-diagonal contribution is

$$e^{ab/\lambda} \sum_{m>0} \left[\left(\frac{b}{a} \right)^{m/2} + \left(\frac{a}{b} \right)^{m/2} \right] (-\iota)^m J_m(2\iota\sqrt{ab}) - e^a - e^b + 2 \tag{A10}$$

Next, we use the definition of the generating function of the Bessel functions (GRADSHTEYN and RYZHIK (2007), 8.511)

$$e^{\frac{1}{2}(t-t^{-1})z} = J_0(z) + \sum_{m>0} (t^m + (-t)^{-m}) J_m(z) \tag{A11}$$

which turns the off-diagonal contribution into

$$e^{ab/\lambda} \left(e^{\sqrt{ab}(\sqrt{\frac{b}{a}} + \sqrt{\frac{a}{b}})} - J_0(2\iota\sqrt{ab}) \right) - e^a - e^b + 2 \tag{A12}$$

Combining the diagonal and off-diagonal contributions and substituting a and b , we find for the integrand in Eq. (A6)

$$e^{ab/\lambda + \sqrt{ab}(\sqrt{\frac{b}{a}} + \sqrt{\frac{a}{b}})} - e^a - e^b + 1 = e^{\lambda e^{-2z-\Delta\tau} - \lambda e^{-z} - \lambda e^{-z-\Delta\tau}} - e^{-\lambda e^{-z}} - e^{-\lambda e^{-z-\Delta\tau}} + 1 \quad (\text{A13})$$

The auto-correlation of x_0 is therefore given by

$$\begin{aligned} \langle x_0(\tau)x_0(\tau + \Delta\tau) \rangle &= \frac{e^{-\lambda}}{Ns} \int_0^\infty dz \left(e^{\lambda e^{-2z-\Delta\tau}} e^{-\lambda(e^{-z} + e^{-z-\Delta\tau})} - e^{-\lambda e^{-z}} - e^{-\lambda e^{-z-\Delta\tau}} + 1 \right) \\ &= \frac{e^{-\lambda}}{Ns} \int_0^1 \frac{d\theta}{\theta} \left(e^{\lambda\theta^2 e^{-\Delta\tau} - \lambda\theta(1+e^{-\Delta\tau})} - e^{-\lambda\theta} - e^{-\lambda\theta e^{-\Delta\tau}} + 1 \right) \end{aligned} \quad (\text{A14})$$

Auto-correlation of the mean fitness

The autocorrelation function of the mean is defined as

$$\begin{aligned} \langle \bar{\delta k}(\tau) \bar{\delta k}(\tau + \Delta\tau) \rangle &= \sum_{i,j>0} ij \langle a_i(\tau) a_j(\tau + \Delta\tau) \rangle \\ &= \partial_\mu \partial_\nu \frac{1}{Ns} \int_0^\infty dz \sum_{i,j>0} \mu^i \nu^j e^{-z(i+j)} \lambda^{i+j} e^{-j\Delta\tau} \sum_{k=0}^{\min(i,j)} \frac{(-1)^{i+j} \lambda^{-k}}{(j-k)!(i-k)!k!} \end{aligned} \quad (\text{A15})$$

where the last line is to be evaluated at $\nu = \mu = 1$. Hence the problem is reduced to the one already solved with $a = -\mu\lambda e^{-z}$ and $b = -\nu\lambda e^{-z-\Delta\tau}$. We find

$$\langle \bar{\delta k}(\tau) \bar{\delta k}(\tau + \Delta\tau) \rangle = \frac{\lambda e^\lambda}{Ns} \int_0^1 d\theta e^{-\Delta\tau} e^{\lambda\theta^2 e^{-\Delta\tau} - \lambda(1+e^{-\Delta\tau})\theta} (\theta + \lambda\theta (\theta e^{-\Delta\tau} - 1) (\theta - 1)) \quad (\text{A16})$$

Cross-correlation of x_0 and the mean fitness

When calculating the cross-correlation between x_0 and the mean fitness we have to distinguish the cases where x_0 precedes the mean fitness and vice-versa. Otherwise, the calculation proceeds almost unchanged from the cases discussed above.

$$\begin{aligned} \langle x_0(\tau) \bar{\delta k}(\tau + \Delta\tau) \rangle &= \sum_{i,j>0} j \langle a_i(\tau) a_j(\tau + \Delta\tau) \rangle \\ &= \partial_\nu \frac{1}{Ns} \int_0^\infty dz \sum_{i,j>0} a^i b^j \sum_{k=0}^{\min(i,j)} \frac{(-1)^{i+j} \lambda^{-k}}{(j-k)!(i-k)!k!} \end{aligned} \quad (\text{A17})$$

with $a = -\lambda e^{-z}$, $b = -\nu\lambda e^{-z-\Delta\tau}$ if $\Delta\tau > 0$ and $a = -\lambda e^{-z+\Delta\tau}$, $b = -\nu\lambda e^{-z}$ if $\Delta\tau < 0$. The result is

$$\langle \delta x_0(\tau) \bar{\delta k}(\tau + \Delta\tau) \rangle = \frac{\lambda}{Ns} \begin{cases} e^{-\Delta\tau} \int_0^1 d\theta \left((\theta - 1) e^{e^{-\Delta\tau} \lambda \theta^2 - \lambda(1+e^{-\Delta\tau})\theta} + e^{-e^{-\Delta\tau} \lambda \theta} \right) & \Delta\tau > 0 \\ \int_0^1 d\theta \left((e^{\Delta\tau} \theta - 1) e^{e^{\Delta\tau} \lambda \theta^2 - \lambda(1+e^{\Delta\tau})\theta} + e^{-\lambda \theta} \right) & \Delta\tau < 0 \end{cases} \quad (\text{A18})$$

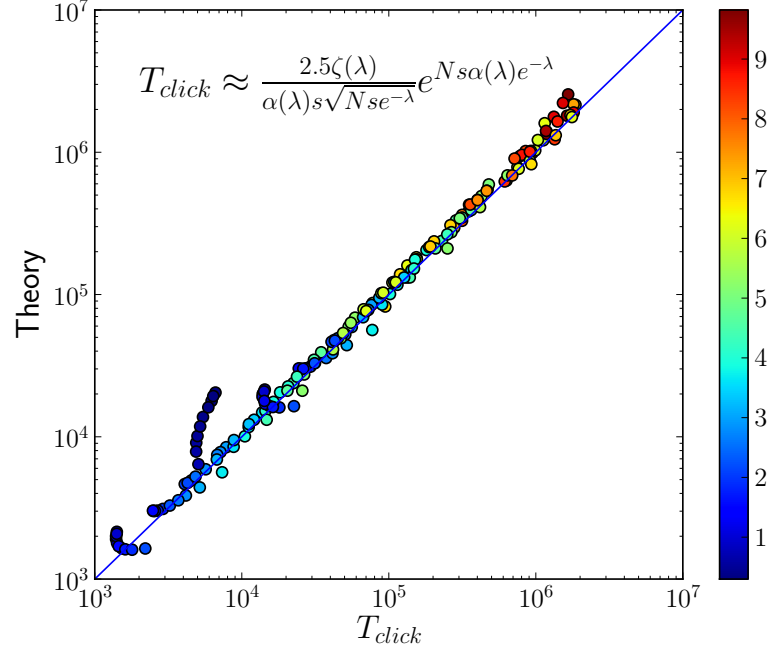


FIG. 6 The approximation of the mean time between clicks of the ratchet is accurate over a large range of parameters if $Ns\alpha(\lambda)e^{-\lambda}$ is large compared to one. $Ns\alpha(\lambda)e^{-\lambda}$ determines whether the clicks of the ratchet are far apart compared to the relaxation time of the distribution and is indicated as the color of the data points. The condition $Ns\alpha(\lambda)e^{-\lambda} > 1$ is violated for the fastest clicks shown, resulting in the deviation of the dark blue points.